

#2

OIPE

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/029,120

DATE: 01/17/2002
TIME: 17:49:50

```
1 <110> APPLICANT: Swanson, Ronald V.
              Feldman, Robert A.
              Schleper, Christa
      4 <120> TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
      5 <130> FILE REFERENCE: DCORP.002A
      6 <140> CURRENT APPLICATION NUMBER: 10/029,120
     7 <141> CURRENT FILING DATE: 2001-12-21
     8 <150> PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020
W--> 9 <151> PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-29
     10 <160> NUMBER OF SEQ ID NOS: 123
     11 <170> SOFTWARE: FastSEQ for Windows Version 3.0
     13 <210> SEO ID NO: 1
     14 <211> LENGTH: 32998
     15 <212> TYPE: DNA
     16 <213> ORGANISM: Cenarchaeum symbiosum
     17 <220> FEATURE:
     18 <221> NAME/KEY: CDS
     19 <222> LOCATION: (7604)...(8908)
                                                          ENTERED
     20 <221> NAME/KEY: CDS
     21 <222> LOCATION: (8961)...(9767)
     22 <221> NAME/KEY: CDS
     23 <222> LOCATION: (10545)...(10922)
     24 <221> NAME/KEY: CDS
     25 <222> LOCATION: (13944)...(14612)
     26 <221> NAME/KEY: CDS
     27 <222> LOCATION: (18638)...(20149)
     28 <221> NAME/KEY: CDS
     29 <222> LOCATION: (20554)...(20955)
     30 <221> NAME/KEY: CDS
     31 <222> LOCATION: (20956)...(21834)
     32 <221> NAME/KEY: CDS
     33 <222> LOCATION: (25151)...(26377)
     34 <221> NAME/KEY: CDS
     35 <222> LOCATION: (27535)...(28002)
     36 <221> NAME/KEY: CDS
     37 <222> LOCATION: (28065)...(29483)
     38 <400> SEQUENCE: 1
                                                                                     60
              gatccttgac ctctgcgctt attgcagcca tggactgacc ggccgtgcgg ggctaaataa
              agetgaggeg cegeetgeag getetgetea geegttgatt atacagtaet egeactegea
                                                                                     120
     40
              ggtcttgctt tcgtcatctt ttgccctgca cttggcgtgc tggcccgact ggcactggac
                                                                                    180
     41
              gcagatctcg tctatgttgt ctatcgaggt gtactttttg gactggtcaa agccgaagcc
                                                                                     240
     42
              gccgtccttt ggtccaacca tgaagcggtc agggccgtta tgccttaatt atctttgccg
                                                                                     300
     43
     44
             teggggeggg geogettetg eeggegggag eeceettga ggeegeteee eggetegttg
                                                                                     360
```

RAW SEQUENCE LISTING DATE: 01/17/2002 PATENT APPLICATION: US/10/029,120 TIME: 17:49:50

							400
45					ctatgattgc		420
46					tgcccatgcg		480
47					ggacagacag		540
48					cacccccac		600
49					ctgggcgggg		660
50					tgagaacaga		720
51	caactgcccg	atcgtttttc	aggtacctac	tgggcttttt	ggagttcgtt	gatcaaaagc	780
52	ggtacaggta	ttctatacat	gccagtcttg	gctggaaaaa	taaattgaag	atcggcggat	840
53	cctatctacg	agcgcctgcc	tgcttttcac	ttgactagcc	ggagtacttc	gtctgcaatt	900
54	tctgtatagc	taggcattta	tgtagttgag	atacatgtcc	gcgggatccg	tttcatagtc	960
·55					aatgcctgat		1020
56					gtttcagctt		1080
57					tttccatctt		1140
58	gtgtttgccg	tggaattacc	taccggggga	acataaaaaa	atgagtcata	aacagcgcac	1200
59					gtgcacagca		1260
60	ccccgcaatc	gtcttatcca	acgacctgcc	ctatgaaaac	tccagacgga	tctcttccgg	1320
61	gcatccagta	cccatgtatg	tgagaattct	gactatctta	ccgtggtgtt	ggcagtgtct	1380
62	caacgacggt	attaccgagg	tgattcatga	taattttgtt	gactggagta	tactgaaaca	1440
63	tggtatttgc	cgcatgaaat	actgacagca	ccagatcttg	aaattcttgt	tcatcttcca	1500
64	gatcagttac	ttttaagcca	atccttttac	attcttctct	cgatatgcgc	cttccatgac	1560
65	tgtgatattt	tttagaggaa	gacattattt	ctgatatttt	ttttgacttt	tttaccgctg	1620
66					tacaagttcc		1680
67	tctggctgct	aatgcatttt	tgaattagcc	cgggaggata	ttgccctaac	agtggaagcc	1740
68	atgcgccgag	cctccccggg	tgtttttctg	acaccgactg	tgcttcttga	aactcgctaa	1800
69	ttagaaactg	tgcagacatt	atgtgcatgc	cggtcctggt	tggaataata	aattgggggt	1860
70	cqqtqqqacc	tatcgatgag	tqtttaccca	ttaccaggca	attcgatgag	catgcaagca	1920
71					gacattttta		1980
72					ccccggactg		2040
73					aaatccatac		2100
74					tatccagttg		2160
75					atagtcgtga		2220
76					gtcgtttgct		2280
77					tcataccata		2340
78	tatgacttgc	cttttttctt	catcaatttc	tcatattcct	catgcagttc	gagcaggatg	2400
79					gcttttggtc		2460
80					tgattttaat		2520
81					acaggcagag		2580
82	gctggggacc	catgaaataa	accccaaca	gtgcacagca	tccgctgggg	gctcaataaa	2640
83	aaaatgagtc	atcatgcata	gtctctatgt	aaatggctga	accggtgttt	tggtcgatta	2700
84	gtaaaggctg	gctcaccact	caccaaaact	tgtgggatac	accaccttcc	tatcaacqca	2760
85	atcttcttct	gcgaaccttc	atccgaagaa	ggaatatctt	gtctcgggat	aggattcgtg	2820
86	cttagatgct	ttcagcactt	agectagatg	gcttagctgc	ccggcctgcc	ctqtcqqaca	2880
87					ctaagagcga		2940
88					cgacggtcta		3000
89	atgttccctt	ttaataggcg	agcagectea	cccttaaccc	ctgctgcagg	accaggatag	3060
90					gagctctcgc		3120
91	cctattatcc	ctagggtaat	ttttctgtca	cetecaaace	ccaatagtgg	gcacacgaag	3180
92					tggaaatcca		3240
93					gaactaaact		3300
<i>)</i>	ceceggeee	Lycocholic	agoggacted	Lyucocycci	Janusaaase		

RAW SEQUENCE LISTING DATE: 01/17/2002 PATENT APPLICATION: US/10/029,120 TIME: 17:49:50

			·				
94	ttgatatctt	ttcaaagggg	tgccgcccca	gccgaactgc	ccacctgcac	atgtccccgg	3360
95					tacatcggcg		3420
96					cactccctgc		3480
97					gatggaagat		3540
98	tcgtccacct	tatgtggctt	caccgggttg	taggcgggga	cagtggggct	ctcgttgttc	3600
99	cattcatgca	cgtcggaact	tacccgacaa	ggcatttggc	taccttaaga	gagtcagagt	3660
100	tactcccgg	c gttaaccggt	ccttagctcg	gttgaaccca	ı agttttagat	accggcaccg	3720
101	gccaggatt	c agcgactata	catacccttt	cgggctagca	gtcgcctgtg	tttttattaa	3780
102	acagtcgaa	a cccccttgtc	actgcaacct	gctgccgcca	ttcctcatga	cagctgcagg	3840
103	catccctta	t acctaagcta	caggactaat	ttgccgaatt	ccctcgccat	acggtatacc	3900
104	cgtagcacc	t tagtttacta	aaccagcgca	cctgtgtcgg	atctgggtac	gaacttgcag	3960
105	tttqctaqc	c gcacggtctt	tcatggtctc	ctggaatcgg	gaaaactctg	ctaacgcaaa	4020
106					cccaggccct		4080
107	-		-		catgcggttc		4140
108					actctgttga		4200
109					accettgege		4260
110					. ctgcaataga		4320
111	caggacgtc	a ccaccctact	togtoccaat	. cactacqcca	acctaccacg	gtgcacctat	4380
112	cacggtgca	c gtctggagta	teggtactet	getttageee	cgtccgtttt	tataacaccc	4440
113	teacteage	a ggtaagttgt	tacacacttt	ttgaaggata	gctacttctg	agettacete	4500
114					. tagcagaaat		4560
115					tacgtcacac		4620
116					atggtgagga		4680
117					acatagcacg		4740
118					tttgacccct		4800
119					agtgggcttt		4860
120	catcttgct	c aggaatagat	cgactggctt	ctageettae	cgccatgact	taacqcactt	4920
121					ccttcggcta		4980
122	ggcttaacc	t coccatoaca	gcaaget.ccc	taacccatat	ttcgagacgg	aacqcatqac	5040
123					aacctccggt		5100
124					tgcacttttc		5160
125					atcggtcttg		5220
126					ccaaggacaa		5280
127					ccctctaagc		5340
128					caaaacacca		5400
129	acattacca	c caactatte	atttaaacto	tttccttttc	gatcgcctct	acttgggaaa	5460
130	tototatta	e gycagailea	toataataat	aagatgette	aattcccacg	attcaacctc	5520
131					gggatcaacg		5580
132					ctcctcaage		5640
133					tacacgacta		5700
134	tastagasa	t greerracae	ggggacacc	agecacatac	tacaccactt	acataataa	5760
	ttacagag	c ecceayyyya	gggggeeget	ttannaga	tttctaagga	gegeggegea	5820
135							5880
136	ccgcaggtt	o occuaeggto	acculgulac	agenciacet	ttgtcgctta	aggggggtg	5940
137	yalaacycc	a accayacyto	accidected	. aayuaaauli	caatgaaacg	tagggeggeg	6000
138	Lytycaagg	a gcagggacgt	. alleactycg	tastasatat	gcgcggttac	gaggattacc	6060
139					ggtagcgttt		6120
140					agcccgcgtg		6180
141					ccgcattaac		6240
142	ccgctaatt	e gececaetge	ceeggagage	aatygtygca	actagaggca	aggatetege	0240

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/029,120

DATE: 01/17/2002
TIME: 17:49:50

143	tcgttacctg	acttaacagg	acatctcacg	gcacgagctg	gcgacggcca	tgcaccacct	6300
144	ctcagcttgt	ctggtagagt	cttcagcttg	accttcacac	tgctgtctct	ccgggtaaga	6360
145	tttctaacat	tgactccaat	tgaaccgcag	gcttcacccc	ttgtggtgct	cccccgccaa	6420
146	ttcctttaag	tatcatactt	gcgtacgtac	ttcccaggcg	gcaaacttaa	cggcttccct	6480
147	geggeactge	actggctctt	acqccaatqc	atcactgagt	ttgcattgtt	tacagctggg	6540
148	actacccggg	tatctaatcc	ggtttgctcc	cccagctttc	atccctcacc	gtcggacgtg	6600
149	ttctagtaga	ccaccttcac	cacagggggt	catcgataga	tcagaggatt	ttaccccttc	6660
150	ctaccgagta	ccgtctacct	ctcccactcc	ctagccgtgc	agtatttccg	gcagcctatg	6720
151	cattgaggg	atagatttaa	ccgaaaactt	acacggcagg	ctacggatgc	tttaggccca	6780
152	ataatcctcc	tgaccacttg	aggtgctggt	tttaccgcgg	cggctgacac	cagaacttgc	6840
153	ccacccctta	ttcaccaata	gttttaagac	cggtaaaaga	tttctttagc	agaaaacact	6900
154	cggattaacc	ttatcatact	ttcqcacatt	gcaaagtttt	ctcgcctgct	gcgccccata	6960
155	gaacctaaat	ccgtgtctca	gtacccatct	ccgggcctct	cctctcagag	cccgtatctg	7020
156	ttatageett	ggtgggccat	tacctcacca	acaagctgat	agaccgcagt	cccatcctac	7080
157	ggcgataaat	catttgggcc	acaaaccatt	ccaggcatag	tggcctatcg	gatattattc	7140
158	teagttteec	gaggttatcc	ccgtccatag	gttagattga	ctacgtgtta	ctgagccgtc	7200
159	taccttatat	toctacaato	actcgcatgg	cttagtatca	atccgatagc	agtcaggtcc	7260
160	ggcaggatca	accggattca	taattggatt	atttttttt	tgttaagtac	gcttgtactt	7320
161	ttggaattga	acagaatgca	cataatcttc	acatctcaga	tatgaccctt	cgatcatacc	7380
162	ctcattctgt	gtgcgtaact	ggaggccagc	gaatcacaat	atggtacaat	accatgcatt	7440
163	categeaage	accactetta	cgtcacgtac	gatcggatcg	gcccgtccat	gggcatataa	7500
164	accategeeg	atttccqccc	ccqqcaqccc	cgatcagggg	ccggatctgc	ctgtatgatg	7560
165	gegateegee	ctgattaaat	tatqqqqqqa	geggeetget	gccgtggatc	tggaacgcga	7620
166	gtacagggca	aagaccggcg	gctcggcccg	gatctttgcc	aggtcgaaaa	agtaccacgt	7680
167	caacaaaatc	agccacaaca	taaggttcta	cgagccgtat	ccgtttgtga	caaggtccgc	7740
168	gagcggcaag	cacctcqtcq	acgtggacgg	gaacaagtat	gtagactact	ggatggggca	7800
169	ctagaaccta	atactggggc	acgcgccggc	gccagtcagg	tcggcagtag	aggggcagct	7860
170	teaccacaac	tggatccacg	ggaccgtcaa	cgagcagacg	atgaatctct	cggagataat	7920
171	acgcggcgcg	gtaagcgtgg	cagaaaagac	aaggtacgtc	acgtcgggga	cggaggccgt	7980
172	catgtatgcg	gcaaggctgg	cgcgcgcgca	tacgggcaga	aaaataatag	caaaggcgga	8040
173	caacaactaa	cacqqqtacq	cgtcggggct	gctcaagtcg	gtcaactggc	cgtatgatgt	8100
174	geeegagage	agagaactca	tcgacgaaga	gcactctata	tccattccgt	acaacgatct	8160
175	tgaaggttcc	ctggatgttc	ttgggcgcgc	aggcgacgac	ttggcatgcg	tgataatcga	8220
176	accactacta	adcadcadca	gctgcatacc	ggcggatgag	gactatctgc	gcggcataca	8280
177	ggagtttgtg	cattcaaggg	gegegetget	tgtcctcgac	gagatagtga	cagggttccg	8340
178	gtttaggttt	ggctgcgcgt	atgctgcagc	agggctggac	cccgatatag	tggcgctcgg	8400
179	caagatagtc	gggggcggat	tccccatagg	ggtgatatgc	ggcaaggacg	aggtgatgga	8460 8520
180	aatctccaac	actatatcgc	atgcaaagtc	cgacagggcg	tacatcggcg	gcggcacatt	
181	ctctgcaaac	cccqccacqa	tgacagcggg	cgcggcagcg	ctcggggagc	tcaaaaagag	8580
182	aaagggcaca	atatacccga	ggataaactc	catgggggac	gacgcaaggg	acaagetete	8640
183	aaagatattt	gggaacaggg	tatccgtgac	cggaaggggc	tcgctgttca	tgactcactt	8700
184	tgttcaagat	ggcgccggca	gggtctcaaa	tgctgcagat	gcggcagcct	gcgatgttga	8760
185	actactacac	aggtaccacc	tggacatgat	cacccgggac	ggcatattct	ttetgeeggg	8820
186	caagetgggg	gccatatcgg	cggcgcactc	aaaggccgac	ctcaagacca	tgtattccgc	8880
187	atcagagege	tttgcagaag	gcctatgagg	tatagcgccg	gaggaaactt	tgattatacg	8940 9000
188	aacatactac	cccqqqqccc	atgatactct	tcggcaagag	cgaccccgcc	gagctggtgc	9000
189	accagacaga	cctcctqtqc	agcaagaacc	agttcagggc	ggcaataggc	ctgtacggga	9120
190	aaatcctcaa	ggacgacccg	cagaacaggg	gcgtcctgca	caaaaagggg	ctggcccaga	9120
191	acagggcaaa	aaagtactct	gatgcgatca	cgtgctttga	ccggctgctc	gagcttgaca	9100

RAW SEQUENCE LISTING DATE: 01/17/2002 PATENT APPLICATION: US/10/029,120 TIME: 17:49:50

192	acaaggacgc	gcccgcgtac	aacaacaagg	ccatagccca	ggccgagctc	ggagacacgg	9240
193	catccgcgct	ggaaaactac	ggcagggcca	tcgaggccga	cccgcggtac	gcgccggcgc	9300
194	gcttcaacag	ggccgtgctg	ctcgacaggc	tgggcgagca	tgaggaggcg	ctgccggacc	9360
195	tcgacagggc	agccgagctg	gaccgacgca	agccgaaccc	gaggttctac	aaggggatag	9420
196	tgctcggcaa	gatgggcagg	cacgaagagg	cgctggcctg	cttcaagggc	gtgtgcaaga	9480
197	ggcatcccgg	ccacgccgac	tcacagttcc	acgtggggat	agagcttacc	gagcttggca	9540
198	ggcacgccga	ggccctcggg	gagcttgcat	cactgcccgc	ggagcaccgc	gagaacgcca	9600
199	atgtattgta	tgccagggcg	cgcagcctct	cgggccttgg	cagggaggac	gaatccatag	9660
200	cqcacctqca	aaaggcggcc	aaaaaagatt	ccaagacgat	aaaaagtgg	gcccgcgcag	9720
201	aaaaggcctt	tgacggaata	cgggacgatc	ccggttcaaa	aagatagccg	gctagaggat	9780
202	ctttttctt	gccgcgtcaa	tccgcatcat	gcggaccttt	tttttgggcc	ccacaagtcg	9840
203	cgattcatag	actggtacat	agaccacctc	caccgccttt	gcggcaaact	cctcccgcag	9900
204	gtcgcgcatg	ccqtcaqqcq	ggggcccgcg	cagcttctct	tttagttttg	agagcgcctc	9960
205	ttctgtctcc	acctcqqqqc	tccgcacatt	ctctgacgca	tcgagtatcc	tccgcgggta	10020
206	cggctccacc	acaccaaacc	ccqtcttqta	gggaaagtcc	gtctcgccgc	cgtgccggtc	10080
207	aaggcacatc	atcccttctq	attccqcaaa	gacatgctct	tctagctcga	ggtcgaccct	10140
208	attettacca	agecetacea	agagcgtctt	gtgtatgcgc	gacttggacc	ttatgggaaa	10200
209	gacgccgt.cg	cctagcacca	cctcgatcac	gttctggtcc	accttgatcg	ggtgaaccgc	10260
210	ctttctgaaa	aaat.ccgcag	agtacctggc	qqaqacccqq	atcagcgcct	cgtggaccag	10320
211	ctttacagaa	tgcacatgga	catcttctt	ccacaaaacc	ctcataaggg	ccctaaaggc	10380
212	accontette	tttgcctcta	tcatggcccg	agccgactcc	tcagtcatgg	cgttccgcag	10440
213	gaccgccgtc	ctagtettte	cagtcatccc	ctgccgcacc	ccgcataagg	catactatac	10500
214	aacgcaaggc	aaggtaataa	tagcctgccg	tctqtaacqq	ccgtatgagg	tcggagggca	10560
215	gacccagata	catcgaaaag	ttcctaaaga	gggcggacaa	ggcgatagac	aatgcagtcg	10620
215	agcaggggg	caagaggga	gacgagatac	tagatgacgc	agtcgagctc	ggcaagatca	10680
217	ccatagagaga	ggcgcaaaaa	agaagcgatg	tactactcaa	gcaggccgag	cgggagagca	10740
217	acconctcaa	gtcaagggg	accaaaaaac	tcgaaaaggg	catagggggg	gcaaaaaaga	10800
219	taacaaccaa	caagggggg	gcgctagaga	ccctggcaaa	gctcggcgag	ctgagaaagg	10860
220	caaaatcat	aacggagaag	gagtttcgcg	ccaagaaaaa	gaagettete	gcggagatct	10920
221	gagttgaage	cactagacta	tacccgggac	ggctcgataa	aggaggtcac	aaagaggtgg	10980
222	tttataggga	caccatacat	tatcaacctt	gcaggcgagc	toggcatato	tgagagcaac	11040
223	atattocaco	tgacatttcc	cdacddcdca	aagaccaccc	tgcatacgca	cgagggcggg	11100
223	canctactca	tagtgacctc	adacaccaac	agcatgtcaa	tatttgaaaa	gaccggcgga	11160
225	gaggaggag	aatttgcaat	aaaagagaga	gacaggatac	cactaaaaca	gggcagcatc	11220
225	ggcgaggcgg	ctaccaacat	actcacatac	accoccaac	agacggcacc	accetgteec	11280
226	atataggggt	aaactacccc	tcaccatcaa	gaaaggaggc	gtatacatta	tggtatgaat	11340
227	acacagegge	aaaccaccccg	acconcatac	tataaattat	attatttgag	cctctccagt	11400
	atagagaga	ttagaaggtt	acceptione	atccccttgc	ggatcacttc	ccttctagtc	11460
229	attyacayyc	acttattasa	ggtcatcgta	ctcttttcac	tggatacete	gagcaccaca	11520
230	etattagaga	tatagagaaa	gagaaacttt	agaggaatca	tcgagcatac	atccatgcgc	11580
231	atgtteeege	cgcacgggaa	gycaaacttt	attaccacta	ccdacacac	ctcgcgtatg	11640
232	cetatteeeg	astataaata	ttggagatag	tacatocata	catageteat	cccaaaaaat	11700
233	tectectegg	adiciccyta	atacatacta	antanataga	aagggcggc	cccgggggat	11760
234	tecetticga	atatetecte	greeargery	ttaaaactta	ccarcarrta	eccgcggccc	11820
235	tccaccgcct	LLALCATET	ggggccgttt	gagaaagtag	caatcacaaa	gtggctggcc	11880
236	tgctgctcaa	agtacggctt	gentinggae	ttgaggtgt	togaggeata	gtacatctcg	11940
237	cctacgttcc	Legacgagga	angagagaga	gootcoatct	tagaggeatg	gtcgtgcgtg	12000
238	tatggcgtgc	aggcatatee	taataaaa	attagagaga	agaagaaaaa	tgggatccgc	12060
239	ggaaccggtt	aatatetagt	ccalgcege	catacttatt	taaaactacc	cctgtggccg	12120
240	gccccggggc	aggegtgegt	ggalecatge	yatayttatt	caaaaccayy	atgccgatca	12120

VERIFICATION SUMMARY

DATE: 01/17/2002

PATENT APPLICATION: US/10/029,120 TIME: 17:49:51

Input Set : N:\Crf3\RULE60\10029120.raw
Output Set: N:\CRF3\01172002\J029120.raw

L:9 M:256 W: Invalid Numeric Header Field, Wrong Prior FILING DATE:YYYY-MM-DD

L:7288 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:120